

SEQUENCE LISTING **AP20 REGD FCT/770** 13 DEC 2005

<110> Lassen, Soren Flensted

<120> Improved proteases and methods for producing them

<130> 10423.204-US

<160> 53

<170> PatentIn version 3.3

<210> 1

<211> 1062

<212> DNA

<213> Nocardiosis sp. NRRL 18262

<220>

<221> misc\_feature

<222> (1)..(495)

<223> Encodes the pro-region shown in positions -165 to -1 of SEQ ID NO:43.

<220>

<221> misc\_feature

<222> (496)..(1059)

<223> Encodes the mature region shown in positions 1-188 of SEQ ID NO:43.

<400> 1

gctactggag cattacctca gtctcctaca cctgaagcag atgcagtatc gatgcaagaa	60
gcattacaac gtgatcttga tcttacatca gctgaagctg aggaattact tgctgcacaa	120
gatacagcct ttgaagttga tgaagctgcc gctgaagcag ctggtgatgc atatggtggt	180
tcagtattcg atactgaatc actcgaactt actgtactag tgaccgatgc agcagctggt	240
gaagctgttg aagccacagg tgcaggtaca gagctcgtat cttatggtat tgatggatta	300
gatgagatcg tacaagagct taatgcagct gatgccgttc caggtgtagt tggatggtat	360
cctgatgtag caggtgatac tgttgtctta gaagttcttg aaggctctgg agctgatggt	420
tctggacttt tagcagacgc aggagtcgat gcatccgcgg ttgaagtgc cacgtcagat	480
cagcctgaac tctatgccga tatcattgga ggcctagcgt acacaatggg tggtcgctgc	540
agcgtaggat ttgcagccac aaatgcagct ggacaacctg gcttcgtgac agctggacat	600
tgcggccgcg tcggtacaca ggttactatc ggcaatggaa gaggtgtctt tgagcaaagc	660
gtatttcccg ggaatgatgc tgccttcgtt agaggtacgt ccaactttac gcttactaac	720
ttagtatcta gatacaacac tggcggatat gcaactgtag caggtcacia tcaagcacct	780

```

attggctcta gcgtctgccg ctcagggtcg actacaggat ggcattgtgg aaccattcaa      840
gctagagggtc agagcgtgag ctatcctgaa ggtaccgtaa cgaacatgac tcgtacgact      900
gtatgtgcag aaccagggtga ctctggaggt tcatatatca gcggtacgca agcgcaaggc      960
gttacctcag gtggatccgg taactgtagg acaggtggca caacgttcta ccaggaagtg     1020
acaccgatgg tgaactcttg gggagttaga ctccgtacat aa                          1062

```

```

<210> 2
<211> 1143
<212> DNA
<213> Artificial sequence

```

```

<220>
<223> A synthetic 10R gene (10Rsyntax-15) encoding a S2A protease denoted
      "10R" fused by PCR in frame to the signal peptide encoding
      sequence of a heterologous protease, Savinase.

```

```

<400> 2
atgaagaaac cgttggggaa aattgtcgca agcaccgcac tactcatttc tgttgctttt      60
agttcatcga tcgcatcggc tgctactgga gcattacctc agtctcctac acctgaagca     120
gatgcagtat cgatgcaaga agcattacaa cgtgatcttg atcttacatc agctgaagct     180
gaggaattac ttgctgcaca agatacagcc tttgaagttg atgaagctgc cgctgaagca     240
gctgggtgatg catatggtgg ttcagtattc gatactgaat cactcgaact tactgtacta     300
gtgaccgatg cagcagctgt tgaagctgtt gaagccacag gtgcaggtag agagctcgta     360
tcttatggta ttgatggatt agatgagatc gtacaagagc ttaatgcagc tgatgccgtt     420
ccagggtgtag ttggatggta tcctgatgta gcaggtgata ctggtgtctt agaagttctt     480
gaaggctctg gagctgatgt ttctggactt ttagcagacg caggagtcga tgcacccgcy     540
gttgaagtga ccacgtcaga tcagcctgaa ctctatgccg atatcattgg aggcctagcg     600
tacacaatgg gtggctcgct cagcgtagga tttgcagcca caaatgcagc tggacaacct     660
ggcttcgtga cagctggaca ttgcggccgc gtcggtacac aggttactat cggcaatgga     720
agagggtgtct ttgagcaaag cgtatttccc gggaatgatg ctgccttcgt tagaggtagc     780
tccaacttta cgcttactaa cttagtatct agatacaaca ctggcggata tgcaactgta     840
gcaggtcaca atcaagcacc tattggctct agcgtctgcc gctcagggtc gactacagga     900
tggcattgtg gaaccattca agctagaggt cagagcgtga gctatcctga aggtaccgta     960
acgaacatga ctcgtacgac tgtatgtgca gaaccagggt actctggagg ttcatatatc    1020
agcggtagcg aagcgcaagg cgttacctca ggtggatccg gtaactgtag gacagggtggc    1080

```

acaacgttct accaggaagt gacaccgatg gtgaactctt ggggagttag actccgtaca 1140

taa 1143

<210> 3

<211> 8

<212> PRT

<213> Artificial sequence

<220>

<223> C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 3

Gln Ser His Val Gln Ser Ala Pro

1

5

<210> 4

<211> 24

<212> DNA

<213> Artificial sequence

<220>

<223> Polynucleotide encoding a C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 4

caatcgcgatg ttcaatccgc tcca

24

<210> 5

<211> 4

<212> PRT

<213> Artificial sequence

<220>

<223> C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 5

Gln Ser Ala Pro

1

<210> 6

<211> 12

<212> DNA

<213> Artificial sequence

<220>

<223> Polynucleotide encoding a C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 6  
caatcggctc ct

12

<210> 7  
<211> 2  
<212> PRT  
<213> Artificial sequence

<220>  
<223> C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 7

Gln Pro  
1

<210> 8  
<211> 6  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Polynucleotide encoding a C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 8  
caacca

6

<210> 9  
<211> 1  
<212> PRT  
<213> Artificial sequence

<220>  
<223> C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 9

Pro  
1

<210> 10  
<211> 3  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Polynucleotide encoding a C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 10  
 cca 3

<210> 11  
 <211> 45  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Primer #252639

<400> 11  
 catgtgcatg tgggtaccgc aacgttcgca gatgctgctg aagag 45

<210> 12  
 <211> 44  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Primer #251992

<400> 12  
 catgtgcatg tggctgaccg attatggagc ggattgaaca tgcg 44

<210> 13  
 <211> 44  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Primer #179541

<400> 13  
 gcgttgagac gcgcggccgc gagcgccgtt tggctgaatg atac 44

<210> 14  
 <211> 43  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Primer #179542

<400> 14  
 gcgttgagac agctcgagca gggaaaaatg gaaccgcttt ttc 43

<210> 15  
 <211> 64  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Primer #179539  
  
 <400> 15  
 ccatttgatc agaattcact ggccgctcgtt ttacaacccat tgcggaaaat agtcataggc 60  
 atcc 64  
  
 <210> 16  
 <211> 60  
 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 <223> Primer #179540  
  
 <400> 16  
 ggatccagat ctggtacccg ggtctagagt cgacgcggcg gttcgcgtcc ggacagcaca 60  
  
 <210> 17  
 <211> 37  
 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 <223> Primer #179154  
  
 <400> 17  
 gttgtaaaac gacggccagt gaattctgat caaatgg 37  
  
 <210> 18  
 <211> 37  
 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 <223> Primer #179153  
  
 <400> 18  
 ccgctcgcac actagacacg ggtacctgat ctagatc 37  
  
 <210> 19  
 <211> 22  
 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 <223> Primer #317  
  
 <400> 19  
 tggcgcaatc ggtaccatgg gg 22

<210> 20  
 <211> 40  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Primer #139 NotI

<400> 20  
 catgtgcatg cggccgcatt aacgcgttgc cgcttctgcg 40

<210> 21  
 <211> 7443  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Sequence of plasmid pMB1508

<400> 21  
 tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60  
 cagcttgtct gtaagcggat gccggggagca gacaagcccg tcagggcgcg tcagcgggtg 120  
 ttggcgggtg tcggggctgg cttactatg cggcatcaga gcagattgta ctgagagtgc 180  
 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240  
 attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat 300  
 tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggg 360  
 tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cgataaaaagt gctttttttg 420  
 ttgcaattga agaattatta atgttaagct taattaaaga taatatcttt gaattgtaac 480  
 gccctcaaa agtaagaact acaaaaaaag aatacgttat atagaaatat gtttgaacct 540  
 tcttcagatt acaaataat tcggacggac tctacctcaa atgcttatct aactatagaa 600  
 tgacatacaa gcacaacctt gaaaatttga aaatataact accaatgaac ttgttcatgt 660  
 gaattatcgc tgtatttaat tttctcaatt caatatataa tatgccaata cattgttaca 720  
 agtagaaatt aagacaccct tgatagcctt actataccta acatgatgta gtattaaatg 780  
 aatatgtaaa tatatttatg ataagaagcg acttatttat aatcattaca tatttttcta 840  
 ttggaatgat taagattcca atagaatagt gtataaatta tttatcttga aaggagggat 900  
 gcctaaaaac gaagaacatt aaaaacatat atttgcaccg tctaattggat ttatgaaaaa 960  
 tcattttatc agtttgaaaa ttatgtatta tggagctctg aaaaaaagga gaggataaag 1020  
 aatgaagaaa ccgttgggga aaattgtcgc aagcaccgca ctactcattt ctgttgcttt 1080

tagttcatcg atcgcatcgg ctgctgaaga agcaaaagaa aaatatttaa ttggctttaa	1140
tgagcaggaa gctgtcagtg agtttgtaga acaagtagag gcaaatagacg aggtcgccat	1200
tctctctgag gaagaggaag tcgaaattga attgcttcat gaatttgaaa cgattcctgt	1260
tttatccggt gagttaagcc cagaagatgt ggacgcgctt gaactcgatc cagcgatttc	1320
ttatattgaa gaggatgcag aagtaacgac aatggcgcaa tcggtaccat ggggtatata	1380
aacgcgttaa tccgcggata tatagcggcc gcagatctgg gaccaataat aatgactaga	1440
gaagaaagaa tgaagattgt tcatgaaatt aaggaacgaa tattggataa agtgggatat	1500
ttttaaaata tatatttatg ttacagtaat attgactttt aaaaaaggat tgattctaata	1560
gaagaaagca gacaagtaag cctcctaaat tcactttaga taaaaattta ggaggcatat	1620
caaatgaact ttaataaaat tgatttagac aattggaaga gaaaagagat atttaatacat	1680
tatttgaacc aacaaacgac ttttagtata accacagaaa ttgatattag tgttttatac	1740
cgaaacataa aacaagaagg atataaattt taccctgcat ttattttctt agtgacaagg	1800
gtgataaact caaatagcgc ttttagaact gggtacaata gcgacggaga gttaggttat	1860
tgggataagt tagagccact ttatacaatt tttgatgggtg tatctaaaac attctctgggt	1920
atttggactc ctgtaaagaa tgacttcaaa gagttttatg atttatacct ttctgatgta	1980
gagaaatata atgggtcggg gaaattgttt cccaaaacac ctatacctga aaatgctttt	2040
tctctttcta ttattccatg gacttcattt actgggttta acttaaataat caataataat	2100
agtaattacc ttctacccat tattacagca ggaaaattca ttaataaagg taattcaata	2160
tatttaccgc tatctttaca ggtacatcat tctgtttgtg atgggttatca tgcaggattg	2220
tttatgaact ctattcagga attgtcagat aggcctaata actggctttt ataatatgag	2280
ataatgccga ctgtactttt tacagtcggt tttctaacga tacattaata ggtacgaaaa	2340
agcaactttt tttgcgctta aaaccagtca taccaataac ttaagggtaa ctacgctcgc	2400
cggaaagagc gaaaatgcct cacatttgtg ccacctaata aggagcgatt tacatatgag	2460
ttatgcagtt tgtagaatgc aaaaagtga atcagctgga ctaaaagggg ccgcagagta	2520
gaatggaaaa ggggatcggg aaacaagtat ataggaggag acctatttat ggcttcagaa	2580
aaagacgcag gaaaacagtc agcagtaaag cttgttccat tgcttattac tgtcgctgtg	2640
ggactaatca tctggtttat tcccgtccg tccggacttg aacctaaagc ttggcatttg	2700
tttgcgattt ttgtcgcaac aattatcggc tttatctcca agcccttgcc aatgggtgca	2760
attgcaattt ttgcattggc gggtactgca ctaactggaa cactatcaat tgaggatata	2820



ttaagcggat tcggaataa gaccatttgg cttatcggtta tcgcattctt tatttcccgg	2880
ggatttatca aaaccggtct cgggtgcgaga atttcgtatg tattcggtca gaaattcgga	2940
aaaaaaaccc ttggactttc ttatttactg ctattcagtg atttaatact ttcacctgct	3000
attccaagta atacggcgcg tgcaggaggc atttatatttc ctattatcag atcattatcc	3060
gaaacattcg gatcaagccc ggcaaatgga acagagagaa aaatcgggtgc attcttatta	3120
aaaaccggtt ttcaggggaa tctgatcaca tctgctatgt tcctgacagc gatggcggcg	3180
aaccgctga ttgccaaagt ggcccatgat gtcgcagggg tggacttaac atggacaagc	3240
tgggcaattg ccgcgattgt accgggactt gtaagcttaa tcatcacgcc gcttgtgatt	3300
tacaaactgt atccgccgga aatcaaagaa acaccgatg cggcgaaaat cgcaacagaa	3360
aaactgaaag aaatgggacc gttcaaaaaa tcggagcttt ccatgggtat cgtgtttctt	3420
ttggtgcttg tgctgtggat ttttggcggc agcttcaaca tcgacgctac cacaaccgca	3480
ttgatcgggt tggccgttct cttattatca caagttctga cttgggatga tatcaagaaa	3540
gaacagggcg cttgggatac gctcacttgg tttgcggcgc ttgtcatgct cgccaacttc	3600
ttgaatgaat taggcatggt gtcttggttc agtaatgcca tgaaatcatc cgtatcaggg	3660
ttctcttga ttgtggcatt catcatttta attgttgtgt attattactc tcactatttc	3720
tttgcaagtg cgacagccca catcagtgcg atgtattcag catttttggc tgtcgtcgtg	3780
gcagcggcg caccgccgct tttagcagcg ctgagcctcg cgttcatcag caacctgttc	3840
gggtcaacga ctactacgg ttctggagcg gctccggtct tcttcggagc aggctacatc	3900
ccgcaaggca aatggtggtc catcggattt atcctgtcga ttgttcatat catcgtatgg	3960
cttgtgatcg gcggattatg gtggaaagta ctaggaatat ggtagaaaga aaaaggcaga	4020
cgcggtctgc ctttttttat tttcactcct tcgtaagaaa atggattttg aaaaatgaga	4080
aaattccctg tgaaaaatgg tatgatctag gtagaaagga cggctggtgc tgtggtgaaa	4140
aagcggttcc atttttccct gcaaacaaaa ataatggggc tgattgcggc tctgctggtc	4200
tttgtcattg gtgtgctgac cattacgtta gccgttcagc atacacaggg agaacggaga	4260
caggcagagc agctggcggg tcaaaccggc agaaccattt cctatatgcc gccggttaaa	4320
gagctcattg agagaaaaga cggacatgcg gctcagacgc aagaggtcat tgaacaaatg	4380
aaagaacaga ctggtgcgtt tgccatttat gttttgaacg aaaaaggaga cattcgcagc	4440
gcctctggaa aaagcggatt aaagaaactg gagcgcagca gagaaatttt gtttggcggg	4500

tcgcatgttt	ctgaaacaaa	agcggatgga	cgaagagtga	tcagagggag	cgcgccgatt	4560
ataaaagaac	agaagggata	cagccaagtg	atcggcagcg	tgtctgttga	ttttctgcaa	4620
acggagacag	agcaaagcat	caaaaagcat	ttgagaaatt	tgagtgtgat	tgctgtgctt	4680
gtactgctgc	tcggatttat	tggcgccgcc	gtgctggcga	aaagcatcag	aaaggatacg	4740
ctcgggcttg	aaccgcatga	gatcgcggct	ctatatcgtg	agaggaacgc	aatgcttttc	4800
gcgattcgag	aagggttat	tgccaccaat	cgtgaaggcg	tcgtcaccat	gatgaacgta	4860
tcggcgcccg	agatgctgaa	gctgcccag	cctgtgatcc	atcttcctat	agatgacgtc	4920
atgccgggag	cagggtgat	gtctgtgctt	gaaaaaggag	aatgctgcc	gaaccaggaa	4980
gtaagcgtca	acgatcaagt	gtttattatc	aatacgaag	tgatgaatca	aggcgggcag	5040
gcgtatggga	ttgtcgtcag	cttcaggag	aaaacagagc	tgaagaagct	gatcgacaca	5100
ttgacagagg	ttcgcaaata	ttcagaggat	ctcaggcgcc	agactcatga	attttcaa	5160
aagctttatg	cgatttttag	gctgcgtcga	cctgcaggca	tgcaagcttg	gcgtaatcat	5220
ggtcatagct	gtttcctgtg	tgaaattgtt	atccgctcac	aattccacac	aacatacgag	5280
ccggaagcat	aaagtgtaaa	gcctgggggtg	cctaattgagt	gagctaactc	acattaattg	5340
cgttgcgctc	actgcccgtc	ttccagtcgg	gaaacctgtc	gtgccagctg	cattaatgaa	5400
tcggccaacg	cgcggggaga	ggcggtttgc	gtattgggcg	ctcttcgct	tcctcgctca	5460
ctgactcgct	gcgctcggtc	gttcggctgc	ggcgagcgg	atcagctcac	tcaaaggcgg	5520
taatacggtt	atccacagaa	tcaggggata	acgcaggaaa	gaacatgtga	gcaaaaggcc	5580
agcaaaaggc	caggaaccgt	aaaaaggccg	cgttgctggc	gtttttccat	aggctccgcc	5640
cccctgacga	gcatcacaaa	aatcgacgct	caagtcagag	gtggcgaaac	ccgacaggac	5700
tataaagata	ccaggcgttt	ccccctggaa	gtccctcgt	gcgctctcct	gttccgaccc	5760
tgccgcttac	cggatacctg	tccgcctttc	tcccttcggg	aagcgtggcg	ctttctcata	5820
gctcacgctg	taggtatctc	agttcgggtg	aggtcgttcg	ctccaagctg	ggctgtgtgc	5880
acgaaccccc	cgttcagccc	gaccgctgcg	ccttatccgg	taactatcgt	cttgagtcca	5940
acccggtaag	acacgactta	tcgccactgg	cagcagccac	tggtaacagg	attagcagag	6000
cgaggtatgt	aggcggtgct	acagagttct	tgaagtgggtg	gcctaactac	ggctacacta	6060
gaaggacagt	atttggatc	tgcgctctgc	tgaagccagt	taccttcgga	aaaagagttg	6120
gtagctcttg	atccggcaaa	caaaccaccg	ctggtagcgg	tggttttttt	gtttgcaagc	6180
agcagattac	gcgcagaaaa	aaaggatctc	aagaagatcc	tttgatcttt	tctacgggggt	6240

ctgacgctca	gtggaacgaa	aactcacggt	aagggatgtt	ggcatgaga	ttatcaaaaa	6300
ggatcttcac	ctagatcctt	ttaaattaaa	aatgaagttt	taaatcaatc	taaagtatat	6360
atgagtaaac	ttggtctgac	agttaccaat	gcttaatcag	tgaggcacct	atctcagcga	6420
tctgtctatt	tcgttcatcc	atagttgcct	gactccccgt	cgtgtagata	actacgatac	6480
gggagggctt	accatctggc	cccagtgtg	caatgatacc	gcgagacca	cgctcaccgg	6540
caccggattt	atcagcaata	aaccagccag	ccggaagggc	cgagcgaga	agtggctctg	6600
caactttatc	cgctccatc	cagtctatta	attggtgccg	ggaagctaga	gtaagtagtt	6660
cgccagttaa	tagtttgcc	aacgttggtg	ccattgctac	aggcatcgtg	gtgtcacgct	6720
cgtcgtttgg	tatggcttca	ttcagctccg	gttcccaacg	atcaaggcga	gttacatgat	6780
cccccatgtt	gtgcaaaaa	gcggttagct	ccttcggtcc	tccgatcgtt	gtcagaagta	6840
agttggccgc	agtgttatca	ctcatgggta	tggcagcact	gcataattct	cttactgtca	6900
tgccatccgt	aagatgcttt	tctgtgactg	gtgagtactc	aaccaagtca	ttctgagaat	6960
agtgtatgcg	gcgaccgagt	tgctcttgcc	cggcgtcaat	acgggataat	accgcgccac	7020
atagcagaac	tttaaaagtg	ctcatcattg	gaaaacgttc	ttcggggcga	aaactctcaa	7080
ggatcttacc	gctgttgaga	tccagttcga	tgtaaccac	tcgtgcaccc	aactgatctt	7140
cagcatcttt	tactttcacc	agcgtttctg	ggtgagcaaa	aacaggaagg	caaatgccg	7200
caaaaaagg	aataaggcg	acacggaaat	gttgaatact	catactcttc	ctttttcaat	7260
attattgaag	catttatcag	ggttattgtc	tcatgagcgg	atacatattt	gaatgtattt	7320
agaaaaataa	acaaatagg	gttccgcgca	catttccccg	aaaagtgcc	cctgacgtct	7380
aagaaacat	tattatcatg	acattaacct	ataaaaatag	gcgtatcacg	aggcccttct	7440
gtc						7443

<210> 22

<211> 5718

<212> DNA

<213> Artificial sequence

<220>

<223> Sequence of MB1510 genomic integration region

<400> 22

gagcgccgtt	tggctgaatg	atacaacagt	ctcacttcct	tactgcgtct	ggttgcaaaa	60
------------	------------	------------	------------	------------	------------	----

acgaagaagc	aaggattccc	ctcgcttctc	atttgtccta	tttattatac	acttttttaa	120
------------	------------	------------	------------	------------	------------	-----

gcacatcttt ggcgcttggt tcactagact tgatgcctct gaatcttgtc caagtgtcac	180
ggtccgcac atagacttgt ccatttttca ccgctttgag atttttccag agcgggttcg	240
ttttccactc atctacaatg gttttgcctt cgttggtgta gatgaacaaa atatcaggat	300
cgattttgct caattgctca aggctgacct cttgataggc gttatctgac ttcacagcgt	360
gtgtaaagcc tagcatttta aagattttctc cgtcatagga tgatgatgta tgaagctgga	420
aggaatccgc tcttgcaacg ccgagaacga tgttgcggtt ttcattcttc ggaagtccg	480
cttttagatc gttgatgact tttttgtgct cggcaagctt ttcttttctt tcatcttctt	540
tatttaatgc tttagcaatg gtcgtaaagc tgtcgatcgt ttcgcatat gtcgcttcac	600
ggctttttta ttcaatcgtc ggggcgattt ttttcagctg ttataaatg tttttatggc	660
gctcagcgtc agcgatgatt aaatcaggct tcaaggaact gatgacctca agattgggtt	720
cgctgcgtgt gcctacagat gtgtaatcaa tggagctgcc gacaagcttt ttaatcatat	780
cttttttggt gtcattctgc atgccaccg gcgtaatgcc gagattgtga acggcatcca	840
agaatgaaag ctcaagcaca accacccgct taggtgtgcc gcttactgtc gtttttctt	900
cttcgcatg gatcactctg gaatccttag actcgctttt gccgcttcg ttgttattct	960
ggcttgatga acagccggat acaatgaggc aggcgagcaa taaaacactc atgatggcaa	1020
tcaacttggt agaatagggt cgcattgcat tcttctttt ttcagattta gtaatgagaa	1080
tcattatcac atgtaacact ataatagcat ggcttatcat gtcaatattt ttttagtaaa	1140
gaaagctgcg tttttactgc tttctcatga aagcatcatc agacacaaat aagtggatg	1200
cagcgttacc gtgtcttcga gacaaaaacg catgggcgtt ggcttttagag gtttcgaaca	1260
tatcagcagt gacataagga aggagagtgc tgagataacc ggacaatttc ttttctattt	1320
catctgttag tgcaaattca atgtcgccga tattcatgat aatcgagaaa acaaagtcga	1380
tatcgatatg aaaatgttcc tcggcaaaaa ccgcaagctc gtgaattcct ggtgaacatc	1440
cggcacgctt atggaaaatc tgtttgacta aatcactcac aatccaagca ttgtattgct	1500
gttctggtga aaagtattgc attagacata cctcctgctc gtacggataa aggcagcgtt	1560
tcatggtcgt gtgctccgtg cagcggcttc tccttaattt tgatttttct gaaaatagg	1620
cccgttcta tcactttacc atggacggaa aacaaatagc tactaccatt cctcctgttt	1680
ttctcttcaa tgttctggaa tctgtttcag gtacagacga tcgggtatga aagaaatata	1740
gaaaacatga aggaggaata tcgacatgaa accagttgta aaagagtata caaatgacga	1800
acagctcatg aaagatgtag aggaattgca gaaaatgggt gttgcgaaag aggatgtata	1860

cgtcttagct	cacgacgatg	acagaacgga	acgcctggct	gacaacacga	acgccaacac	1920
gatcggagcc	aaagaaacag	gtttcaagca	cgcggtggga	aatatcttca	ataaaaaagg	1980
agacgagctc	cgcaataaaa	ttcacgaaat	cggtttttct	gaagatgaag	ccgctcaatt	2040
tgaaaaacgc	ttagatgaag	gaaaagtgct	tctctttgtg	acagataacg	aaaaagtga	2100
agcttgggca	taaagcaagg	aaaaaaccaa	aaggccaatg	tcggcctttt	ggtttttttg	2160
cggtctttgc	ggtgggattt	tgcaaatgc	cgcaatagga	tagcggaaca	tttccggttc	2220
tgaatgtccc	tcaatttgct	attatatatt	tgtgataaat	tggaataaaa	tctcacaaaa	2280
tagaaaatgg	gggtacatag	tggaatgaaa	aagtgatgtt	agctacggct	ttgttttttag	2340
gattgactcc	agctggcgcg	aacgcagctg	atttaggcca	ccagacgttg	ggatccaatg	2400
atggctgggg	cgcgctactc	accggcacga	caggcggatc	aaaagcatcc	tcctcaaagt	2460
tgtataccgt	cagcaacaga	aaccagcttg	tctcggcatt	aggaaggaa	acgaacacaa	2520
cgccaaaaat	catttatatc	aagggaacga	ttgacatgaa	cgtggatgac	aatctgaagc	2580
cgcttggcct	aaatgactat	aaagatccgg	agtatgattt	ggacaaatat	ttgaaagcct	2640
atgatcctag	cacatggggc	aaaaaagagc	cgtcgggaac	acaagaagaa	gcgagagcac	2700
gctctcagaa	aaacccaaaa	gcacgggtca	tggtggatat	ccctgcaaac	acgacgatcg	2760
tcggttcagg	gactaacgct	aaagtcgtgg	gaggaaactt	ccaaatcaag	agtgataacg	2820
tcattattcg	caacattgaa	ttccaggatg	cctatgacta	tttccgcaa	tggttgtaaa	2880
acgacggcca	gtgaattctg	atcaaattgg	tcagtgagag	cgaagcgaac	acttgatttt	2940
ttaattttct	atcttttata	ggtcattaga	gtatacttat	ttgtcctata	aactattttag	3000
cagcataata	gatttattga	ataggtcatt	taagttgagc	atattagagg	aggaaaatct	3060
tgagagaaata	tttgaagaac	ccgagatcta	gatcaggtag	cgcaacgttc	gcagatgctg	3120
ctgaagagat	tattaaaaag	ctgaaagcaa	aaggctatca	attggtaact	gtatctcagc	3180
ttgaagaagt	gaagaagcag	agaggctatt	gaataaatga	gtagaaagcg	ccatatcggc	3240
gcttttcttt	tggaagaaaa	tatagggaaa	atggtacttg	ttaaaaattc	ggaatattta	3300
tacaatatca	tatgtatcac	attgaaagga	ggggcctgct	gtccagactg	tccgctgtgt	3360
aaaaataagg	aataaagggg	ggttgacatt	attttactga	tatgtataat	ataatttgta	3420
taagaaaatg	gaggggccct	cgaaacgtaa	gatgaaacct	tagataaaaag	tgcttttttt	3480
gttgcaattg	aagaattatt	aatgttaagc	ttaattaaag	ataatatctt	tgaattgtaa	3540

cgcccctcaa aagtaagaac tacaaaaaaa gaatacgtta tatagaaata tgtttgaacc	3600
ttcttcagat tacaaatata ttcggacgga ctctacctca aatgcttata taactataga	3660
atgacatata agcacaacct tgaaaatttg aaaatataac taccaatgaa cttgttcatg	3720
tgaattatcg ctgtatttaa ttttctcaat tcaatatata atatgccaat acattgttac	3780
aagtagaaat taagacaccc ttgatagcct tactatacct aacatgatgt agtattaaat	3840
gaatatgtaa atatatttat gataagaagc gacttattta taatcattac atatttttct	3900
attggaatga ttaagattcc aatagaatag tgtataaatt atttatcttg aaaggaggga	3960
tgcctaaaaa cgaagaacat taaaaacata tatttgcacc gtctaatagga tttatgaaaa	4020
atcattttat cagtttgaaa attatgtatt atggagctct gaaaaaaagg agaggataaa	4080
gagaaaaggg gatcggaaaa caagtatata ggaggagacc tatttatggc ttcagaaaaa	4140
gacgcaggaa aacagtcagc agtaaagctt gttccattgc ttattactgt cgctgtggga	4200
ctaatacctt ggtttattcc cgctccgtcc ggacttgaac ctaaagcttg gcatttgttt	4260
gcgatttttg tcgcaacaat tatcggcttt atctccaagc ccttgccaat ggggtgcaatt	4320
gcaatttttg cattggcggc tactgcacta actggaacac tatcaattga ggatacatta	4380
agcggattcg ggaataagac catttggctt atcgttatcg cattctttat ttcccgggga	4440
tttatcaaaa ccggtctcgg tgcgagaatt tcgtatgtat tcgttcagaa attcggaaaa	4500
aaaacccttg gactttctta ttactgcta ttcagtgatt taatactttc acctgctatt	4560
ccaagtaata cggcgcgtgc aggaggcatt atatttccta ttatcagatc attatccgaa	4620
acattcggat caagcccggc aaatggaaca gagagaaaaa tcggtgcatt cttattaa	4680
accggttttc aggggaatct gatcacatct gctatgttcc tgacagcgat ggcggcgaac	4740
ccgctgattg ccaagctggc ccatgatgtc gcaggggtgg acttaacatg gacaagctgg	4800
gcaattgccg cgattgtacc gggacttgta agcttaatca tcacgccgct tgtgatttac	4860
aaactgtatc cgccggaaat caaagaaaca ccggatgcgg cgaaaatcgc aacagaaaaa	4920
ctgaaagaaa tgggaccgtt caaaaaatcg gagctttcca tggttatcgt gtttcttttg	4980
gtgcttggtc tgtggatttt tggcggcagc ttcaacatcg acgctaccac aaccgcattg	5040
atcggtttgg ccgttctctt attatcacia gttctgactt gggatgatat caagaaagaa	5100
cagggcgctt gggatacgct cacttggttt gcggcgcttg tcatgctcgc caacttcttg	5160
aatgaattag gcatggtgtc ttggttcagt aatgccatga aatcatccgt atcagggttc	5220
tcttggattg tggcattcat cattttaatt gttgtgtatt attactctca ctatttcttt	5280

gcaagtgcga cagccacat cagtgcgatg tattcagcat ttttggtgt cgtcgtggca 5340  
gcgggcgcac cgccgctttt agcagcgctg agcctcgcgt tcacagcaa cctgttcggg 5400  
tcaacgactc actacggttc tggagcggct ccggtcttct tcggagcagg ctacatcccg 5460  
caaggcaaat ggtggtccat cggatttata ctgtcgattg ttcatatcat cgtatggctt 5520  
gtgatcggcg gattatggtg gaaagtacta ggaatatggt agaaagaaaa aggcagacgc 5580  
ggtctgcctt tttttatttt cactccttcg taagaaaatg gattttgaaa aatgagaaaa 5640  
ttccctgtga aaaatggtat gatctaggta gaaaggacgg ctggtgctgt ggtgaaaaag 5700  
cggttccatt tttccctg 5718

<210> 23  
<211> 27  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer 1605

<400> 23  
gacggccagt gaattcgata aaagtgc 27

<210> 24  
<211> 42  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer 1606

<220>  
<221> misc\_feature  
<222> (13)..(13)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (16)..(16)  
<223> n is a, c, g, or t

<400> 24  
ccagatctct atnktntgt acggagtcta actccccaag ag 42

<210> 25  
<211> 1112  
<212> DNA  
<213> Nocardiosis dassonvillei DSM 43235

<400> 25  
gcttttagtt catcgatcgc atcggctgct ccggcccccg tccccagac cccgctcgcc 60  
gacgacagcg ccgccagcat gaccgaggcg ctcaagcgcg acctcgacct cacctcggcc 120  
gaggccgagg agcttctctc ggcgaggaa gccgccatcg agaccgacgc cgaggccacc 180  
gaggccgcg gcgaggccta cggcggtca ctgttcgaca ccgagaccct cgaactcacc 240  
gtgctggtca ccgacgcctc cgccgtcgag gcggtcgagg ccaccggagc ccaggccacc 300  
gtcgtctccc acggcaccga gggcctgacc gaggtcgtgg aggacctcaa cggcgccgag 360  
gttcccagaga gcgtcctcgg ctggtacccg gacgtggaga gcgacaccgt cgtggtcgag 420  
gtgctggagg gtcggacgc cgacgtcgcc gccctgctcg ccgacgccgg tgtggactcc 480  
tcctcgggtcc ggggtggagga ggccgaggag gcccgcagg tctacgccga catcatcggc 540  
ggcctggcct actacatggg cgcccgctgc tccgtcgggt tcgccgcgac caacagcgcc 600  
ggtcagcccg gtttcgtcac cgccggccac tggggcaccg tcggcaccgg cgtgaccatc 660  
ggcaacggca ccggcacctt ccagaactcg gtcttccccg gcaacgacgc cgccttcgtc 720  
cgcggcacct ccaacttcac cctgaccaac ctggtctcgc gctacaactc cggcggtac 780  
cagtcggtga ccggtaccag ccaggccccg gccggctcgg ccgtgtgccg ctccggctcc 840  
accaccggct ggcactgcgg caccatccag gcccgcaacc agaccgtgcg ctaccgcgag 900  
ggcaccgtct actcgctcac ccgcaccaac gtgtgcgccg agcccggcga ctccggcggt 960  
tcgttcatct ccggctcgca ggcccagggc gtcacctccg gcggctccgg caactgctcc 1020  
gtcggcgga cgacctacta ccaggaggtc accccgatga tcaactcctg ggggtgcagg 1080  
atccggacct aatcgcatgt tcaatccgct cc 1112

<210> 26  
<211> 48  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer 1423

<400> 26  
gcttttagtt catcgatcgc atcggctgct ccggcccccg tccccag 48

<210> 27  
<211> 45  
<212> DNA  
<213> Artificial sequence



<220>

<223> Primer 1475

<400> 27

ggagcggatt gaacatgcga ttaggtccgg atcctgacac cccag

45

<210> 28

<211> 354

<212> PRT

<213> Nocardiosis dassonvillei DSM 43235

<220>

<221> PROPEP

<222> (1)..(166)

<220>

<221> mat\_peptide

<222> (167)..(354)

<400> 28

Ala Pro Ala Pro Val Pro Gln Thr Pro Val Ala Asp Asp Ser Ala  
-165 -160 -155

Ala Ser Met Thr Glu Ala Leu Lys Arg Asp Leu Asp Leu Thr Ser  
-150 -145 -140

Ala Glu Ala Glu Glu Leu Leu Ser Ala Gln Glu Ala Ala Ile Glu  
-135 -130 -125

Thr Asp Ala Glu Ala Thr Glu Ala Ala Gly Glu Ala Tyr Gly Gly  
-120 -115 -110

Ser Leu Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp  
-105 -100 -95

Ala Ser Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gln Ala Thr Val  
-90 -85 -80 -75

Val Ser His Gly Thr Glu Gly Leu Thr Glu Val Val Glu Asp Leu Asn  
-70 -65 -60

Gly Ala Glu Val Pro Glu Ser Val Leu Gly Trp Tyr Pro Asp Val Glu  
-55 -50 -45

Ser Asp Thr Val Val Val Glu Val Leu Glu Gly Ser Asp Ala Asp Val

-40                      -35                      -30  
 Ala Ala Leu Leu Ala Asp Ala Gly Val Asp Ser Ser Ser Val Arg Val  
     -25                      -20                      -15  
 Glu Glu Ala Glu Glu Ala Pro Gln Val Tyr Ala Asp Ile Ile Gly Gly  
 -10                      -5                      -1 1                      5  
 Leu Ala Tyr Tyr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr  
           10                      15                      20  
 Asn Ser Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Thr  
           25                      30                      35  
 Val Gly Thr Gly Val Thr Ile Gly Asn Gly Thr Gly Thr Phe Gln Asn  
     40                      45                      50  
 Ser Val Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn  
 55                      60                      65                      70  
 Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly Tyr Gln  
           75                      80                      85  
 Ser Val Thr Gly Thr Ser Gln Ala Pro Ala Gly Ser Ala Val Cys Arg  
           90                      95                      100  
 Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn  
     105                      110                      115  
 Gln-Thr Val Arg Tyr Pro Gln Gly Thr Val Tyr Ser Leu Thr Arg Thr  
     120                      125                      130  
 Asn Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser Gly  
 135                      140                      145                      150  
 Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Val  
           155                      160                      165  
 Gly Gly Thr Thr Tyr Tyr Gln Glu Val Thr Pro Met Ile Asn Ser Trp  
           170                      175                      180  
 Gly Val Arg Ile Arg Thr  
     185

<210> 29  
 <211> 498  
 <212> DNA  
 <213> Nocardiosis dassonvillei DSM 43235

<400> 29  
 gctccggccc ccgtccccc gacccccgtc gccgacgaca gcgccgccag catgaccgag 60  
 gcgctcaagc gcgacctga cctcacctcg gccgaggccg aggagcttct ctcggcgcag 120  
 gaagccgcca tcgagaccga cgccgaggcc accgaggccg cgggcgaggc ctacggcggc 180  
 tcactgttcg acaccgagac cctcgaactc accgtgctgg tcaccgacgc ctccgccgtc 240  
 gaggcggtcg aggccaccgg agcccaggcc accgtcgtct cccacggcac cgagggcctg 300  
 accgaggtcg tggaggacct caacggcgcc gaggttcccg agagcgtcct cggctggtac 360  
 ccggacgtgg agagcgacac cgtcgtggtc gaggtgctgg agggctccga cgccgacgtc 420  
 gccgccctgc tcgccgacgc cgggtgtggac tcctcctcgg tccgggtgga ggaggccgag 480  
 gagggccgc aggtctac 498

<210> 30  
 <211> 166  
 <212> PRT  
 <213> Nocardiosis dassonvillei DSM 43235

<400> 30

Ala	Pro	Ala	Pro	Val	Pro	Gln	Thr	Pro	Val	Ala	Asp	Asp	Ser	Ala	Ala
1				5					10					15	
Ser	Met	Thr	Glu	Ala	Leu	Lys	Arg	Asp	Leu	Asp	Leu	Thr	Ser	Ala	Glu
			20					25					30		
Ala	Glu	Glu	Leu	Leu	Ser	Ala	Gln	Glu	Ala	Ala	Ile	Glu	Thr	Asp	Ala
			35				40					45			
Glu	Ala	Thr	Glu	Ala	Ala	Gly	Glu	Ala	Tyr	Gly	Gly	Ser	Leu	Phe	Asp
	50					55					60				
Thr	Glu	Thr	Leu	Glu	Leu	Thr	Val	Leu	Val	Thr	Asp	Ala	Ser	Ala	Val
65					70					75				80	
Glu	Ala	Val	Glu	Ala	Thr	Gly	Ala	Gln	Ala	Thr	Val	Val	Ser	His	Gly
			85					90						95	

Thr Glu Gly Leu Thr Glu Val Val Glu Asp Leu Asn Gly Ala Glu Val  
100 105 110

Pro Glu Ser Val Leu Gly Trp Tyr Pro Asp Val Glu Ser Asp Thr Val  
115 120 125

Val Val Glu Val Leu Glu Gly Ser Asp Ala Asp Val Ala Ala Leu Leu  
130 135 140

Ala Asp Ala Gly Val Asp Ser Ser Ser Val Arg Val Glu Glu Ala Glu  
145 150 155 160

Glu Ala Pro Gln Val Tyr  
165

<210> 31  
<211> 1146  
<212> DNA  
<213> Artificial sequence

<220>  
<223> The DNA sequence coding for the pro-region of SEQ ID NO: 29 fused  
in frame to A1918L2 protease tail-variant encoding gene; whole  
construct: 10R(proA1918L2).

<400> 31  
atgaagaaac cggtggggaa aattgtcgca agcaccgcac tactcatttc tgttgctttt 60  
agttcatcga tcgcatcggc tgctccggcc cccgtccccc agacccccgt cgccgacgac 120  
agcgccgcca gcatgaccga ggcgctcaag cgcgacctcg acctcacctc ggccgaggcc 180  
gaggagcttc tctcggcgca ggaagccgcc atcgagaccg acgccgaggc caccgaggcc 240  
gcggggcgagg cctacggcgg ctcaactgttc gacaccgaga ccctcgaact caccgtgctg 300  
gtcaccgacg cctccgccgt cgaggcggtc gaggccaccg gagcccaggc caccgtcgtc 360  
tcccacggca ccgagggcct gaccgaggtc gtggaggacc tcaacggcgc cgaggttccc 420  
gagagcgtcc tcggctggta cccggacgtg gagagcgaca ccgtcgtggt cgaggtgctg 480  
gagggctccg acgcccacgt cgccgccctg ctcgccgacg ccggtgtgga ctctcctcgc 540  
gtccgggtgg aggaggccga ggaggccccg caggtctatg ccgatatcat tggaggccta 600  
gcgtacacaa tgggtggtcg ctgcagcgta ggatttgag ccacaaatgc agctggacaa 660  
cctggcttcg tgacagctgg acattgcggc cgcgtcggta cacaggttac tatcggcaat 720  
ggaagaggtg tctttgagca aagcgtattt cccgggaatg atgctgcctt cgttagaggt 780

acgtccaact ttacgcttac taacttagta tctagataca acactggcgg atatgcaact	840
gtagcaggtc acaatcaagc acctattggc tctagcgtct gccgctcagg gtcgactaca	900
ggatggcatt gtggaaccat tcaagctaga ggtcagagcg tgagctatcc tgaaggtacc	960
gtaacgaaca tgactcgtac gactgtatgt gcagaaccag gtgactctgg aggttcatat	1020
atcagcggta cgcaagcgca aggcgttacc tcaggtggat ccggttaactg taggacaggt	1080
ggcacaacgt tctaccagga agtgacaccg atggtgaact cttggggagt tagactccgt	1140
acataa	1146

<210> 32

<211> 1068

<212> DNA

<213> Nocardiosis Alba DSM 15647

<400> 32

gcgaccggcc ccctcccca gtccccacc ccggatgaag ccgaggccac caccatggtc	60
gaggccctcc agcgcgacct cggcctgtcc ccctctcagg ccgacgagct cctcgaggcg	120
caggccgagt ccttcgagat cgacgaggcc gccaccgcgg ccgcagccga ctctacggc	180
ggctccatct tcgacaccga cagcctcacc ctgaccgtcc tggtcaccga cgcctccgcc	240
gtcgaggcgg tcgaggccgc cggcgccgag gccaaagggtg tctcgcacgg catggagggc	300
ctggaggaga tcgtcgccga cctgaacgcg gccgacgtc agcccggcgt cgtgggctgg	360
taccccgaca tccactccga cacggtcgtc ctcgaggctc tcgagggtc cggtgccgac	420
gtggactccc tgctcgccga cgccggtgtg gacaccgccg acgtcaagggt ggagagcacc	480
accgagcagc ccgagctgta cgccgacatc atcggcggtc tcgcctacac catgggtggg	540
cgtgtctcgg tcggcttcgc ggccaccaac gcctccggcc agcccgggtt cgtcaccgcc	600
ggccactgcg gcaccgtcgg caccgccgtc agcatcgga acggccaggg cgtcttcgag	660
cgttccgtct tccccggaa cgactccgcc ttcgtccgcg gcacctgaa cttaccctg	720
accaacctgg tcagccgcta caacaccggt ggttacgcga ccgtctccgg ctctcgag	780
gcggcgatcg gctcgcagat ctgccgttcc ggtccacca ccggtggca ctgcggcacc	840
gtccaggccc gcggccagac ggtgagctac cccagggca ccgtgcagaa cctgaccgc	900
accaacgtct gcgccgagcc cggtgactcc ggcggtcct tcactccgg cagccaggcc	960
cagggcgtca cctccggtgg ctccggcaac tgctccttcg gtggcaccac ctactaccag	1020
gaggtcaacc cgatgctgag cagctggggt ctgaccctgc gcacctga	1068

<210> 33  
 <211> 355  
 <212> PRT  
 <213> Nocardiosis Alba DSM 15647

<220>  
 <221> PROPEP  
 <222> (1)..(167)

<220>  
 <221> mat\_peptide  
 <222> (168)..(355)

<400> 33

Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Asp Glu Ala Glu  
 -165 -160 -155

Ala Thr Thr Met Val Glu Ala Leu Gln Arg Asp Leu Gly Leu Ser  
 -150 -145 -140

Pro Ser Gln Ala Asp Glu Leu Leu Glu Ala Gln Ala Glu Ser Phe  
 -135 -130 -125

Glu Ile Asp Glu Ala Ala Thr Ala Ala Ala Ala Asp Ser Tyr Gly  
 -120 -115 -110

Gly Ser Ile Phe Asp Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr  
 -105 -100 -95

Asp Ala Ser Ala Val Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys  
 -90 -85 -80

Val Val Ser His Gly Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu  
 -75 -70 -65 -60

Asn Ala Ala Asp Ala Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile  
 -55 -50 -45

His Ser Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp  
 -40 -35 -30

Val Asp Ser Leu Leu Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys  
 -25 -20 -15

Val Glu Ser Thr Thr Glu Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly  
 -10 -5 -1 1 5

Gly Leu Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala  
 10 15 20

Thr Asn Ala Ser Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly  
 25 30 35

Thr Val Gly Thr Pro Val Ser Ile Gly Asn Gly Gln Gly Val Phe Glu  
 40 45 50

Arg Ser Val Phe Pro Gly Asn Asp Ser Ala Phe Val Arg Gly Thr Ser  
 55 60 65

Asn Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr  
 70 75 80 85

Ala Thr Val Ser Gly Ser Ser Gln Ala Ala Ile Gly Ser Gln Ile Cys  
 90 95 100

Arg Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Val Gln Ala Arg  
 105 110 115

Gly Gln Thr Val Ser Tyr Pro Gln Gly Thr Val Gln Asn Leu Thr Arg  
 120 125 130

Thr Asn Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser  
 135 140 145

Gly Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser  
 150 155 160 165

Phe Gly Gly Thr Thr Tyr Tyr Gln Glu Val Asn Pro Met Leu Ser Ser  
 170 175 180

Trp Gly Leu Thr Leu Arg Thr  
 185

<210> 34  
 <211> 43  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Primer 1421

<400> 34  
 gttcatcgat cgcacggct gcgaccggcc cctccccca gtc 43

<210> 35  
 <211> 31  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Primer 1604

<400> 35  
 gcggtaccta tcaggtgcgc agggtcagac c 31

<210> 36  
 <211> 1062  
 <212> DNA  
 <213> Nocardiosis prasina DSM 15648

<400> 36  
 gccaccggac cgctccccca gtcaccacc ccggaggccg acgccgtctc catgcaggag 60  
 gcgctccagc gcgacctcg cctgaccccg cttgaggccg atgaactgct ggccgcccag 120  
 gacaccgcct tcgaggtcga cgaggccgcg gccgcggccg ccggggacgc ctacggcggc 180  
 tccgtcttcg acaccgagac cctggaactg accgtccttg tcaccgacgc cgcctcggtc 240  
 gaggtgtgag aggccaccgg cgcggtacc gaactcgtct cctacggcat cgagggcctc 300  
 gacgagatca tccaggatct caacgccg ccgacccgtcc ccggcgtggt cggctggtac 360  
 ccggacgtgg cgggtgacac cgtcgtcctg gaggtccttg agggttccgg agccgacgtg 420  
 agcggcctgc tcgccgacgc cggcgtggac gcctcggccg tcgaggtgac cagcagtgcg 480  
 cagcccagac tctacgccga catcatcggc ggtctggcct acaccatggg cggccgctgt 540  
 tcggtcggat tcgcggccac caacgccg ccgtcagccc gattcgtcac cgccggtcac 600  
 tgtggccgcg tgggcaccca ggtgagcatc ggcaacggcc agggcgtctt cgagcagtcc 660  
 atcttcccgg gcaacgacgc cgccttcgtc cgcggcacgt ccaacttcac gctgaccaac 720  
 ctggtcagcc gctacaacac cggcgtgtac gccaccgtcg ccggccacaa ccaggcgccc 780  
 atcggctcct ccgtctgccg ctccggctcc accaccggct ggcactgcgg caccatccag 840  
 gcccgcgcc agtcggtgag ctaccccgag ggcaccgtca ccaacatgac ccggaccacc 900  
 gtgtgcgcc agcccgccga ctccggcggc tcctacatct ccggcaacca ggcccagggc 960



gtcacctccg gcgggtccgg caactgccgc accggcggga ccaccttcta ccaggaggtc 1020

accccatgg tgaactcctg gggcggtccgt ctccggacct aa 1062

<210> 37

<211> 353

<212> PRT

<213> Nocardiosis prasina DSM 15648

<220>

<221> PROPEP

<222> (1)..(165)

<220>

<221> mat\_peptide

<222> (166)..(353)

<400> 37

Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala  
-165 -160 -155

Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro  
-150 -145 -140

Leu Glu Ala Asp Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu  
-135 -130 -125

Val Asp Glu Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly  
-120 -115 -110

Ser Val Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp  
-105 -100 -95 -90

Ala Ala Ser Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu  
-85 -80 -75

Val Ser Tyr Gly Ile Glu Gly Leu Asp Glu Ile Ile Gln Asp Leu Asn  
-70 -65 -60

Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala  
-55 -50 -45

Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val  
-40 -35 -30

Ser Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val  
 -25 -20 -15 -10

Thr Ser Ser Ala Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu  
 -5 -1 1 5

Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn  
 10 15 20

Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val  
 25 30 35

Gly Thr Gln Val Ser Ile Gly Asn Gly Gln Gly Val Phe Glu Gln Ser  
 40 45 50 55

Ile Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe  
 60 65 70

Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr  
 75 80 85

Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser  
 90 95 100

Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln  
 105 110 115

Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr  
 120 125 130 135

Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Asn  
 140 145 150

Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly  
 155 160 165

Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly  
 170 175 180

Val Arg Leu Arg Thr  
 185

<210> 38  
 <211> 43  
 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 <223> Primer 1346  
  
 <400> 38  
 gttcatcgat cgcacgcggct gccaccggac cgctcccca gtc 43  
  
 <210> 39  
 <211> 38  
 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 <223> Primer 1602  
  
 <400> 39  
 gcggatccta ttaggtccgg agacggacgc cccaggag 38  
  
 <210> 40  
 <211> 1062  
 <212> DNA  
 <213> Nocardiosis prasina DSM 15649  
  
 <400> 40  
 gccaccggac cactcccca gtcaccacc ccggaggccg acgccgtctc catgcaggag 60  
 gcgctccagc gcgacctcgg cctgaccccg cttgaggccg atgaactgct ggccgcccag 120  
 gacaccgcct tcgaggtcga cgaggcccg gccgaggccg ccggtgacgc ctacggcggc 180  
 tccgtcttcg acaccgagac cctggaactg accgtcctgg tcaccgactc cgccgcggtc 240  
 gaggcggtgg aggccaccgg cgccgggacc gaactggtct cctacggcat cacgggcctc 300  
 gacgagatcg tcgaggagct caacgccgcc gacgccgttc ccggcggtgg cggctggtac 360  
 ccggacgtcg cgggtgacac cgtcgtgctg gaggtcctgg agggttccgg cgccgacgtg 420  
 ggccggcctgc tcgccgacgc cggcgtggac gcctcggcgg tcgaggtgac caccaccgag 480  
 cagcccagac tgtacgccga catcatcggc ggtctggcct acaccatggg cggccgctgt 540  
 tcggtcggct tcgcggccac caacgccgcc ggtcagcccg ggttcgtcac cgccggtcac 600  
 tgtggcccg tgggcaccca ggtgaccatc ggcaacggcc gggcgctctt cgagcagtcc 660  
 atcttcccgg gcaacgacgc cgccttcgtc cgcggaacgt ccaacttcac gctgaccaac 720  
 ctggtcagcc gctacaacac cggcggctac gccaccgtcg ccggtcacia ccaggcgccc 780  
 atcggctcct ccgtctgccg ctccgggtcc accaccggtt ggcaactgcgg caccatccag 840

gcccgcggcc agtcggtgag ctaccccgag ggcaccgtca ccaacatgac gcggaccacc 900  
gtgtgcgccg agcccggcga ctccggcggc tctacatct cgggcaacca ggcccagggc 960  
gtcacctccg gcggctccgg caactgccgc accggcgggg ccaccttcta ccaggaggtc 1020  
accccatgg tgaactcctg gggcgctccgt ctccggacct aa 1062

<210> 41  
<211> 353  
<212> PRT  
<213> Nocardiosis prasina DSM 15649

<220>  
<221> PROPEP  
<222> (1)..(165)

<220>  
<221> mat\_peptide  
<222> (166)..(353)

<400> 41

Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala  
-165 -160 -155

Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro  
-150 -145 -140

Leu Glu Ala Asp Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu  
-135 -130 -125

Val Asp Glu Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly  
-120 -115 -110

Ser Val Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp  
-105 -100 -95 -90

Ser Ala Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu  
-85 -80 -75

Val Ser Tyr Gly Ile Thr Gly Leu Asp Glu Ile Val Glu Glu Leu Asn  
-70 -65 -60

Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala  
-55 -50 -45

Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val  
 -40 -35 -30

Gly Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val  
 -25 -20 -15 -10

Thr Thr Thr Glu Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu  
 -5 -1 1 5

Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn  
 10 15 20

Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val  
 25 30 35

Gly Thr Gln Val Thr Ile Gly Asn Gly Arg Gly Val Phe Glu Gln Ser  
 40 45 50 55

Ile Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe  
 60 65 70

Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr  
 75 80 85

Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser  
 90 95 100

Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln  
 105 110 115

Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr  
 120 125 130 135

Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Asn  
 140 145 150

Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly  
 155 160 165

Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly  
 170 175 180

Val Arg Leu Arg Thr  
185

<210> 42  
<211> 43  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer 1603

<400> 42  
gttcacgat cgcacggct gccaccggac cactccccca gtc

43

<210> 43  
<211> 353  
<212> PRT  
<213> Nocardiosis sp. NRRL 18262

<220>  
<221> PROPEP  
<222> (1)..(165)

<220>  
<221> mat\_peptide  
<222> (166)..(1059)

<400> 43

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala  
-165 -160 -155

Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser  
-150 -145 -140

Ala Glu Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu  
-135 -130 -125

Val Asp Glu Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly  
-120 -115 -110

Ser Val Phe Asp Thr Glu Ser Leu Glu Leu Thr Val Leu Val Thr Asp  
-105 -100 -95 -90

Ala Ala Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu  
-85 -80 -75

Val Ser Tyr Gly Ile Asp Gly Leu Asp Glu Ile Val Gln Glu Leu Asn

-70	-65	-60
Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala		
-55	-50	-45
Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val		
-40	-35	-30
Ser Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val		
-25	-20	-15
Thr Thr Ser Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu		
-5	-1 1	5
Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn		
10	15	20
Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val		
25	30	35
Gly Thr Gln Val Thr Ile Gly Asn Gly Arg Gly Val Phe Glu Gln Ser		
40	45	50
Val Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe		
60	65	70
Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr		
75	80	85
Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser		
90	95	100
Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln		
105	110	115
Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr		
120	125	130
Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Thr		
140	145	150
Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly		
155	160	165

Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly  
 170 175 180

Val Arg Leu Arg Thr  
 185

<210> 44  
 <211> 1164  
 <212> DNA  
 <213> artificial sequence

<220>  
 <223> Synthetic protease encoding gene

<220>  
 <221> CDS  
 <222> (1)..(1164)  
 <223> Full length protease

<220>  
 <221> sig\_peptide  
 <222> (1)..(81)

<220>  
 <221> misc\_feature  
 <222> (82)..(1164)  
 <223> Propeptide

<220>  
 <221> mat\_peptide  
 <222> (577)..(1164)

<400> 44		
atg aaa aaa ccg ctg gga aaa att gtc gca agc aca gca ctt ctt		45
Met Lys Lys Pro Leu Gly Lys Ile Val Ala Ser Thr Ala Leu Leu		
-190 -185 -180		
att tca gtg gca ttt agc tca tct att gca tca gca gct aca gga		90
Ile Ser Val Ala Phe Ser Ser Ser Ile Ala Ser Ala Ala Thr Gly		
-175 -170 -165		
gca tta ccg cag tct ccg aca ccg gaa gca gat gca gtc tca atg		135
Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val Ser Met		
-160 -155 -150		
caa gaa gca ctg caa aga gat ctt gat ctt aca tca gca gaa gca		180
Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu Ala		
-145 -140 -135		
gaa gaa ctt ctt gct gca caa gat aca gca ttt gaa gtg gat gaa		225
Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu		
-130 -125 -120		



gca gcg gca gaa gca gca gga gat gca tat ggc ggc tca gtt ttt Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly Ser Val Phe -115 -110 -105	270
gat aca gaa tca ctt gaa ctt aca gtt ctt gtt aca gat gca gca gca Asp Thr Glu Ser Leu Glu Leu Thr Val Leu Val Thr Asp Ala Ala Ala -100 -95 -90	318
gtt gaa gca gtt gaa gca aca gga gca gga aca gta ctt gtt tca tat Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Val Leu Val Ser Tyr -85 -80 -75	366
gga att gat ggc ctt gat gaa att gtt caa gaa ctg aat gca gct gat Gly Ile Asp Gly Leu Asp Glu Ile Val Gln Glu Leu Asn Ala Ala Asp -70 -65 -60 -55	414
gct gtt ccg ggc gtt gtt ggc tgg tat ccg gat gtt gct gga gat aca Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly Asp Thr -50 -45 -40	462
gtt gtc ctt gaa gtt ctt gaa gga tca ggc gca gat gtt tca ggc ctg Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Ser Gly Leu -35 -30 -25	510
ctg gca gac gca gga gtc gat gca tca gca gtt gaa gtt aca aca tca Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Thr Ser -20 -15 -10	558
gat caa ccg gaa ctt tat gca gat att att ggc ggc ctg gca tat tat Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu Ala Tyr Tyr -5 -1 1 5 10	606
atg ggc ggc aga tgc agc gtt ggc ttt gca gca aca aat gca tca ggc Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn Ala Ser Gly 15 20 25	654
caa ccg ggc ttt gtt aca gca ggc cat tgc ggc aca gtt ggc aca cca Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Thr Val Gly Thr Pro 30 35 40	702
gtt tca att ggc aat ggc aaa ggc gtt ttt gaa cga agc att ttt ccg Val Ser Ile Gly Asn Gly Lys Gly Val Phe Glu Arg Ser Ile Phe Pro 45 50 55	750
ggc aat gat tca gca ttt gtt aga ggc aca tca aat ttt aca ctt aca Gly Asn Asp Ser Ala Phe Val Arg Gly Thr Ser Asn Phe Thr Leu Thr 60 65 70	798
aat ctg gtt tca aga tat aat tca ggc ggc tat gca aca gtt gca ggc Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly Tyr Ala Thr Val Ala Gly 75 80 85 90	846
cat aat caa gca ccg att ggc tca gca gtt tgc aga tca ggc tca aca His Asn Gln Ala Pro Ile Gly Ser Ala Val Cys Arg Ser Gly Ser Thr 95 100 105	894

aca ggc tgg cat tgc ggc aca att caa gca aga aat caa aca gtt agg 942  
 Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn Gln Thr Val Arg  
 110 115 120

tat ccg caa ggc aca gtt tat agt ctg aca aga aca aca gtt tgt gca 990  
 Tyr Pro Gln Gly Thr Val Tyr Ser Leu Thr Arg Thr Thr Val Cys Ala  
 125 130 135

gaa ccg ggc gat tca ggc ggc tca tat att agc ggc act caa gca caa 1038  
 Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Thr Gln Ala Gln  
 140 145 150

ggc gtt aca tca ggc ggc tca ggc aat tgc agt gct ggc ggc aca aca 1086  
 Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Ala Gly Gly Thr Thr  
 155 160 165 170

tat tac caa gaa gtt aat ccg atg ctt agt tca tgg ggc ctt aca ctt 1134  
 Tyr Tyr Gln Glu Val Asn Pro Met Leu Ser Ser Trp Gly Leu Thr Leu  
 175 180 185

aga aca caa tcg cat gtt caa tcc gct cca 1164  
 Arg Thr Gln Ser His Val Gln Ser Ala Pro  
 190 195

<210> 45  
 <211> 388  
 <212> PRT  
 <213> artificial sequence

<220>  
 <223> Synthetic Construct

<400> 45

Met Lys Lys Pro Leu Gly Lys Ile Val Ala Ser Thr Ala Leu Leu  
 -190 -185 -180

Ile Ser Val Ala Phe Ser Ser Ser Ile Ala Ser Ala Ala Thr Gly  
 -175 -170 -165

Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val Ser Met  
 -160 -155 -150

Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu Ala  
 -145 -140 -135

Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu  
 -130 -125 -120

Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly Ser Val Phe  
 -115 -110 -105

Asp Thr Glu Ser Leu Glu Leu Thr Val Leu Val Thr Asp Ala Ala Ala  
 -100 -95 -90

Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Val Leu Val Ser Tyr  
 -85 -80 -75

Gly Ile Asp Gly Leu Asp Glu Ile Val Gln Glu Leu Asn Ala Ala Asp  
 -70 -65 -60 -55

Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly Asp Thr  
 -50 -45 -40

Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Ser Gly Leu  
 -35 -30 -25

Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Thr Ser  
 -20 -15 -10

Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu Ala Tyr Tyr  
 -5 -1 1 5 10

Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn Ala Ser Gly  
 15 20 25

Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Thr Val Gly Thr Pro  
 30 35 40

Val Ser Ile Gly Asn Gly Lys Gly Val Phe Glu Arg Ser Ile Phe Pro  
 45 50 55

Gly Asn Asp Ser Ala Phe Val Arg Gly Thr Ser Asn Phe Thr Leu Thr  
 60 65 70

Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly Tyr Ala Thr Val Ala Gly  
 75 80 85 90

His Asn Gln Ala Pro Ile Gly Ser Ala Val Cys Arg Ser Gly Ser Thr  
 95 100 105

Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn Gln Thr Val Arg  
 110 115 120

Tyr Pro Gln Gly Thr Val Tyr Ser Leu Thr Arg Thr Thr Val Cys Ala  
125 130 135

Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Thr Gln Ala Gln  
140 145 150

Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Ala Gly Gly Thr Thr  
155 160 165 170

Tyr Tyr Gln Glu Val Asn Pro Met Leu Ser Ser Trp Gly Leu Thr Leu  
175 180 185

Arg Thr Gln Ser His Val Gln Ser Ala Pro  
190 195

<210> 46  
<211> 165  
<212> PRT  
<213> Artificial sequence

<220>  
<223> Shuffled pro-peptide O-2.19

<220>  
<221> PROPEP  
<222> (1)..(165)

<400> 46

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val  
1 5 10 15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu  
20 25 30

Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu  
35 40 45

Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly Ser Val Phe Asp  
50 55 60

Thr Glu Ser Leu Thr Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val  
65 70 75 80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly

	85		90		95										
Met	Glu	Gly	Leu	Glu	Glu	Ile	Val	Ala	Asp	Leu	Asn	Ala	Ala	Asp	Ala
			100					105					110		
Gln	Pro	Gly	Val	Val	Gly	Trp	Tyr	Pro	Asp	Ile	His	Ser	Asp	Thr	Val
			115				120					125			
Val	Leu	Glu	Val	Leu	Glu	Gly	Ser	Gly	Ala	Asp	Val	Asp	Ser	Leu	Leu
	130					135					140				
Ala	Asp	Ala	Gly	Val	Asp	Ala	Ser	Ala	Val	Glu	Val	Thr	Thr	Ser	Asp
145					150					155					160
Gln	Pro	Glu	Leu	Tyr											
				165											

<210> 47  
 <211> 166  
 <212> PRT  
 <213> Artificial sequence  
 <220>  
 <223> Shuffled propeptide G-2.73  
 <220>  
 <221> PROPEP  
 <222> (1)..(166)  
 <400> 47

Ala	Thr	Gly	Ala	Leu	Pro	Gln	Ser	Pro	Thr	Pro	Glu	Ala	Asp	Ala	Val
1				5					10					15	
Ser	Met	Gln	Glu	Ala	Leu	Gln	Arg	Asp	Leu	Asp	Leu	Ser	Ser	Ala	Glu
			20					25					30		
Ala	Glu	Glu	Leu	Leu	Ala	Ala	Gln	Asp	Thr	Ala	Phe	Glu	Val	Asp	Glu
		35					40					45			
Ala	Ala	Ala	Gly	Ala	Ala	Gly	Asp	Ala	Tyr	Gly	Gly	Ser	Val	Phe	Asp
	50					55					60				
Thr	Glu	Thr	Leu	Glu	Leu	Thr	Val	Leu	Val	Thr	Asp	Ala	Ser	Ala	Val
65					70					75					80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly  
85 90 95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala  
100 105 110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val  
115 120 125

Val Val Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu  
130 135 140

Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr  
145 150 155 160

Glu Gln Pro Glu Leu Tyr  
165

<210> 48  
<211> 166  
<212> PRT  
<213> Artificial sequence

<220>  
<223> Shuffled propeptide G-1.43

<220>  
<221> PROPEP  
<222> (1)..(166)

<400> 48

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val  
1 5 10 15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Ser Ser Ser Gln  
20 25 30

Ala Glu Glu Leu Leu Asp Ala Gln Ala Glu Ser Phe Glu Ile Asp Glu  
35 40 45

Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly Ser Ile Phe Asp  
50 55 60

Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val

65                                      70                                      75                                      80  
 Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly  
                                     85                                      90                                      95  
 Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala  
                                     100                                      105                                      110  
 Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val  
                                     115                                      120                                      125  
 Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu  
                                     130                                      135                                      140  
 Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr  
                                     145                                      150                                      155                                      160  
 Glu Gln Pro Glu Leu Tyr  
                                     165  
  
 <210> 49  
 <211> 166  
 <212> PRT  
 <213> Artificial sequence  
  
 <220>  
 <223> Shuffled propeptide G-2.6  
  
 <220>  
 <221> PROPEP  
 <222> (1)..(166)  
  
 <400> 49  
  
 Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val  
 1                                      5                                      10                                      15  
  
 Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu  
                                     20                                      25                                      30  
  
 Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu  
                                     35                                      40                                      45  
  
 Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly Ser Ile Phe Asp  
                                     50                                      55                                      60

Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp Ser Ser Ser Val  
65 70 75 80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly  
85 90 95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala  
100 105 110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val  
115 120 125

Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu  
130 135 140

Ala Gly Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr  
145 150 155 160

Glu Gln Pro Glu Leu Tyr  
165

<210> 50  
<211> 165  
<212> PRT  
<213> Artificial sequence

<220>  
<223> Shuffled propeptide G-2.5

<220>  
<221> PROPEP  
<222> (1)..(165)

<400> 50

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val  
1 5 10 15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro Leu Glu  
20 25 30

Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu  
35 40 45

Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly Ser Val Phe Asp





Glu Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly Ser Ile Phe  
 50 55 60

Asp Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr Asp Ala Ala Ala  
 65 70 75 80

Val Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His  
 85 90 95

Gly Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp  
 100 105 110

Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly Asp Thr  
 115 120 125

Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Tyr Ser Leu  
 130 135 140

Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Pro Ala  
 145 150 155 160

Ala Gln Pro Glu Leu Tyr  
 165

<210> 52  
 <211> 166  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <223> Shuffled propeptide G-1.4

<220>  
 <221> PROPEP  
 <222> (1)..(166)

<400> 52

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val  
 1 5 10 15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Ser Ser Ser Gln  
 20 25 30

Ala Glu Glu Leu Leu Asp Ala Gln Ala Glu Ser Phe Glu Ile Asp Glu

35                      40                      45  
 Ala Ala Ala Ala Ala Ala Ala Asp Ser Tyr Gly Gly Ser Ile Phe Asp  
     50                      55                      60  
 Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val  
 65                      70                      75                      80  
 Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly  
                     85                      90                      95  
 Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala  
                     100                      105                      110  
 Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val  
                     115                      120                      125  
 Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu  
                     130                      135                      140  
 Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr  
 145                      150                      155                      160  
 Glu Gln Pro Glu Leu Tyr  
                     165

<210> 53  
 <211> 166  
 <212> PRT  
 <213> Artificial sequence

<220>  
 <223> Shuffled propeptide G-1.2

<220>  
 <221> PROPEP  
 <222> (1)..(166)

<400> 53

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val  
 1                      5                      10                      15  
 Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu  
                     20                      25                      30

Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu  
 35 40 45

Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly Ser Ile Phe Asp  
 50 55 60

Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp Ser Ser Ser Val  
 65 70 75 80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly  
 85 90 95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala  
 100 105 110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val  
 115 120 125

Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu  
 130 135 140

Ala Gly Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr  
 145 150 155 160

Glu Gln Pro Glu Leu Tyr  
 165